

# On multivariate trace inequalities of Sutter, Berta, and Tomamichel

Marius Lemm

Citation: [Journal of Mathematical Physics](#) **59**, 012204 (2018);

View online: <https://doi.org/10.1063/1.5001009>

View Table of Contents: <http://aip.scitation.org/toc/jmp/59/1>

Published by the [American Institute of Physics](#)

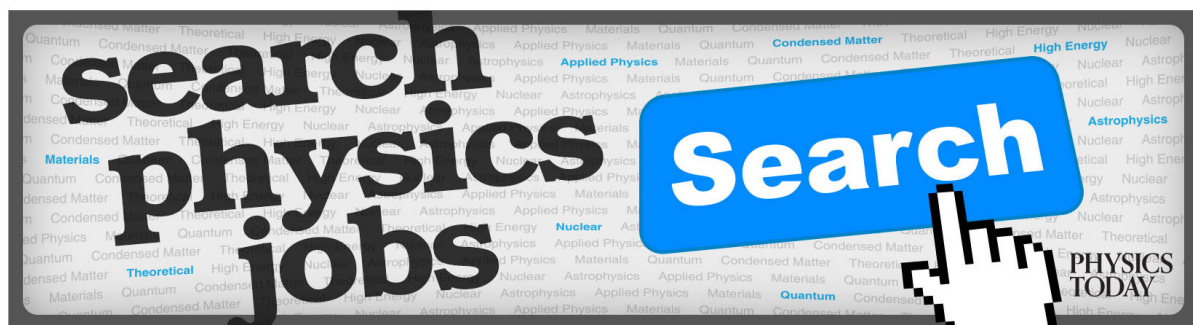
---

## Articles you may be interested in

[Rigorous derivation of nonlinear Dirac equations for wave propagation in honeycomb structures](#)

[Journal of Mathematical Physics](#) **59**, 011509 (2018); 10.1063/1.5021754

---



# On multivariate trace inequalities of Sutter, Berta, and Tomamichel

Marius Lemm<sup>a)</sup>

*Department of Mathematics, California Institute of Technology, Pasadena, CA 91125, USA  
 and Department of Mathematics, Harvard University, Cambridge, Massachusetts 02138, USA*

(Received 21 August 2017; accepted 27 November 2017; published online 23 January 2018)

We consider a family of multivariate trace inequalities recently derived by Sutter, Berta, and Tomamichel. These inequalities generalize the Golden-Thompson inequality and Lieb's triple matrix inequality to an arbitrary number of matrices in a way that features complex matrix powers (i.e., certain unitaries). We show that their inequalities can be rewritten as an  $n$ -matrix generalization of Lieb's original triple matrix inequality. The complex matrix powers are replaced by resolvents and appropriate maximally entangled states. We expect that the technically advantageous properties of resolvents, in particular for perturbation theory, can be of use in applications of the  $n$ -matrix inequalities, e.g., for analyzing the performance of the rotated Petz recovery map in quantum information theory and for removing the unitaries altogether. *Published by AIP Publishing.* <https://doi.org/10.1063/1.5001009>

## I. INTRODUCTION AND MAIN RESULT

Arguably the most famous trace inequality is the Golden-Thompson inequality,<sup>6,19</sup> which asserts that for all positive definite matrices  $A_1$  and  $A_2$ ,

$$\mathrm{Tr}[\exp(\log A_1 + \log A_2)] \leq \mathrm{Tr}[A_1 A_2]. \quad (1.1)$$

In 1973, Lieb proved a generalization of (1.1) to three positive definite matrices  $A_1$ ,  $A_2$ , and  $A_3$ , namely,

$$\mathrm{Tr}[\exp(\log A_1 + \log A_2 + \log A_3)] \leq \mathrm{Tr}\left[A_3 T_{A_2^{-1}}(A_1)\right]. \quad (1.2)$$

The right-hand side features  $T_X(Y)$ , the Fréchet derivative of the matrix logarithm

$$T_X(Y) := \left. \frac{d}{dr} \log(X + rY) \right|_{r=0} = \int_0^\infty \frac{1}{X + \tau} Y \frac{1}{X + \tau} d\tau, \quad (1.3)$$

where  $X$  and  $Y$  are positive definite matrices. Here and in the following, we write  $\tau$  for  $\tau I$ , where  $I$  is an identity matrix of the appropriate dimension. The expression  $T_X(Y)$  can be understood as a non-commutative analog of  $X^{-1}Y$ . Lieb's triple matrix inequality (1.2) has numerous applications, in particular, to entropy inequalities in quantum information theory.

In 2016, Sutter, Berta, and Tomamichel<sup>16</sup> generalized Lieb's triple matrix inequality to the following *multivariate trace inequality*:

$$\mathrm{Tr}\left[\exp\left(\sum_{k=1}^n \log A_k\right)\right] \leq \int_{\mathbb{R}} \mathrm{Tr}[A_n A_{n-1}^{\frac{1+it}{2}} \dots A_2^{\frac{1+it}{2}} A_1 A_2^{\frac{1-it}{2}} \dots A_{n-1}^{\frac{1-it}{2}}] \beta(t) dt. \quad (1.4)$$

Here  $A_1, \dots, A_n$  are positive definite matrices and

$$\beta(t) := \frac{\pi}{2} (1 + \cosh(\pi t))^{-1} \quad (1.5)$$

<sup>a)</sup> [mlemm@caltech.edu](mailto:mlemm@caltech.edu)

is an explicit probability density function. The  $n = 4$  version of (1.4) has applications to recoverability questions in quantum information theory, which we will discuss later on.

The proof of (1.4) is based on powerful complex interpolation techniques.<sup>1,9,15</sup> (In fact, these techniques yield a stronger inequality, namely an  $n$ -matrix version of the Araki-Lieb-Thirring inequality for Schatten norms;<sup>16</sup> see also Ref. 8. The derivation of the  $n$ -matrix inequality (1.4) is discussed following Corollary 3.3 in Ref. 16.) Using the same techniques,<sup>1,9,15</sup> closely related results were proved in Refs. 4 and 22.

We mention that there exist other multivariate trace inequalities generalizing the Golden-Thompson inequality, in particular one proved by Hansen,<sup>7</sup> and we refer to Ref. 16 for a summary of the pertinent literature.

Given these results, it is natural to relate the (*a priori* rather different looking) inequalities (1.2) and (1.4) for  $n = 3$ . In this vein, in Ref. 16 it was observed that

$$\int_{\mathbb{R}} A_2^{\frac{1+i\tau}{2}} A_1 A_2^{\frac{1-i\tau}{2}} \beta(\tau) d\tau = T_{A_2^{-1}}(A_1). \quad (1.6)$$

By diagonalizing  $A_2$ , this is seen to be equivalent to the equality  $\sqrt{xy} \int (y/x)^{it/2} \beta(t) dt = \int_0^\infty (x^{-1} + \tau)^{-1} (y^{-1} + \tau)^{-1} d\tau$ , which holds for all real numbers  $x, y > 0$ . The identity (1.6) implies that when  $n = 3$ , inequality (1.4) is the same as Lieb's triple matrix inequality (1.2) since

$$\int_{\mathbb{R}} \text{Tr}[A_3 A_2^{\frac{1+i\tau}{2}} A_1 A_2^{\frac{1-i\tau}{2}}] \beta(\tau) d\tau = \text{Tr}[A_3 T_{A_2^{-1}}(A_1)]. \quad (1.7)$$

## A. The main result

The main result, Theorem 1.1 below, shows how to generalize the form of Lieb's triple matrix inequality (1.2) to  $n$  matrices.

The generalization is derived from the  $n$ -matrix inequality (1.4), by appropriately rewriting its right-hand side. Specifically, we observe that the identity (1.7) can be generalized to  $n$  matrices by introducing a suitable, non-trivial tensor product structure (Lemma 2.1).

In preparation for the main result, Theorem 1.1, we introduce some notations:

- Given  $x \in \mathbb{R}$ , we write  $\lceil x \rceil$  for the smallest *positive* integer greater or equal to  $x$ . Note that this means  $\lceil 0 \rceil = 1$ .
- We write  $\{\alpha_k\}_{k \geq 2}$  for the *Thue-Morse sequence*. The Thue-Morse sequence is a  $\{0, 1\}$ -valued sequence which can be defined as the substitution sequence associated with the substitution rules  $0 \rightarrow 01$  and  $1 \rightarrow 10$ , with initial value 0.

We use the following equivalent definition: We set  $\alpha_{j+2} := \beta_j$  ( $j \geq 0$ ), where the sequence  $\{\beta_j\}_{j \geq 0}$  is inductively defined by the relations

$$\beta_0 = 0, \quad \beta_{2j} = \beta_j, \quad \beta_{2j+1} = 1 - \beta_j. \quad (1.8)$$

For example, the first four elements of the Thue-Morse sequence are  $(\alpha_2, \alpha_3, \alpha_4, \alpha_5) = (0, 1, 1, 0)$ .

- Let  $\mathcal{H}$  be a  $d$ -dimensional Hilbert space, with a fixed orthonormal basis  $\{|l\rangle\}_{l=1}^d$ . Given  $m \geq 1$ , we write  $|\Omega_m\rangle$  for the maximally entangled state on  $\mathcal{H}^{\otimes m} \otimes \mathcal{H}^{\otimes m}$  associated with this basis, i.e.,

$$|\Omega_m\rangle := \sum_{l_1, \dots, l_m=1}^d |l_1 \otimes \dots \otimes l_m\rangle \otimes |l_1 \otimes \dots \otimes l_m\rangle.$$

We emphasize that  $|\Omega_m\rangle$  is *not* normalized, i.e.,  $\|\Omega_m\| = d^{m/2} \neq 1$ . We denote the corresponding (also non-normalized) projector by

$$P_m = |\Omega_m\rangle \langle \Omega_m|.$$

We refer to Refs. 20 and 23 for background on maximally entangled states. (We emphasize that we use a different normalization convention, however.)

- We write  $\overline{A_k}$  for the complex conjugate of the positive definite matrix  $A_k$ . By definition, the complex conjugate is taken with respect to the fixed  $\{|l\rangle\}_l$  basis from above. (We mention the fact that the matrix  $\overline{A_k}$  is still positive definite and that  $A_k = A_k^T$ .) We write  $C$  for the complex conjugation operator in the  $\{|l\rangle\}_l$  basis. For any matrix  $X$  on  $\mathcal{H}$ , it holds that

$$C^{\alpha_k} X C^{\alpha_k} := \begin{cases} X, & \text{if } \alpha_k = 0, \\ \overline{X}, & \text{if } \alpha_k = 1. \end{cases} \quad (1.9)$$

Our main result is the following alternative, equivalent form of the multivariate trace inequality (1.4) from Ref. 16. We recall the definitions (1.3) of  $T_X(Y)$  and (1.5) of  $\beta(t)$ .

**Theorem 1.1.** *Let  $n \geq 3$  and let  $A_1, \dots, A_n$  be positive definite matrices on a finite-dimensional Hilbert space  $\mathcal{H}$ . Set  $n' := \lceil \log_2(n-2) \rceil$  and  $\rho = 2^{\lceil \log_2(n-2) \rceil} - (n-2)$ . We have the inequality*

$$\begin{aligned} & \text{Tr}_{\mathcal{H}} \left[ \exp \left( \sum_{k=1}^n \log A_k \right) \right] \\ & \leq \text{Tr}_{\mathcal{H}^{\otimes 2^{n'}}} \left[ P_{2^{n'-1}} T_{\bigotimes_{k=2}^{n-1} C^{\alpha_k} A_k^{-1} C^{\alpha_k} \otimes I_{\mathcal{H}}^{\otimes \rho}} \left( A_1 \otimes \overline{A_n} \otimes \bigotimes_{j=0}^{n'-2} P_{2^j} \right) \right]. \end{aligned} \quad (1.10)$$

We may write out the matrix  $T_X(Y)$  appearing in (1.10) as an integral by using (1.3). When we do so, (1.10) reads

$$\begin{aligned} & \text{Tr}_{\mathcal{H}} \left[ \exp \left( \sum_{k=1}^n \log A_k \right) \right] \\ & \leq \int_0^\infty \text{Tr}_{\mathcal{H}^{\otimes 2^{n'}}} \left[ P_{2^{n'-1}} \left( \frac{1}{\bigotimes_{k=2}^{n-1} C^{\alpha_k} A_k^{-1} C^{\alpha_k} + \tau} \otimes I_{\mathcal{H}}^{\otimes \rho} \right) \right. \\ & \quad \left. \left( A_1 \otimes \overline{A_n} \otimes \bigotimes_{j=0}^{n'-2} P_{2^j} \right) \left( \frac{1}{\bigotimes_{k=2}^{n-1} C^{\alpha_k} A_k^{-1} C^{\alpha_k} + \tau} \otimes I_{\mathcal{H}}^{\otimes \rho} \right) \right] d\tau. \end{aligned}$$

*Remark 1.2.* (i) The identity matrix  $I_{\mathcal{H}}^{\otimes \rho}$  is present for dimensional reasons. We note that  $\rho = 0$  iff  $n-2 = 2^m$  for some integer  $m$ .

- (ii) We recall that the first four elements of the Thue-Morse sequence  $\{\alpha_k\}_{k \geq 2}$  are  $(0, 1, 1, 0)$ . Therefore, we have

$$\bigotimes_{k=2}^5 C^{\alpha_k} X_k C^{\alpha_k} = X_2 \otimes \overline{X_3} \otimes \overline{X_4} \otimes X_5,$$

and this already suffices for many cases of interest.

- (iii) If  $n = 3$ , then (1.10) is just Lieb's triple matrix inequality (1.2). Indeed, in this case, we have  $n' = \lceil 0 \rceil = 1$  per our convention and (1.10) says

$$\text{Tr}_{\mathcal{H}} [\log(A_1 + \log A_2 + \log A_3)] \leq \text{Tr}_{\mathcal{H}^{\otimes 2}} \left[ P_1 T_{A_2 \otimes I_{\mathcal{H}}} (A_1 \otimes \overline{A_3}) \right] = \text{Tr}_{\mathcal{H}} [A_3 T_{A_2} (A_1)].$$

The second equality is an elementary property of the maximally entangled state; its general form is recalled in (1.12) below.

- (iv) If  $n = 4$ , then  $n' = 1$  and (1.10) gives the new inequality

$$\text{Tr}_{\mathcal{H}} [\exp(\log A_1 + \log A_2 + \log A_3 + \log A_4)] \leq \text{Tr}_{\mathcal{H}^{\otimes 2}} \left[ P_1 T_{(A_2 \otimes \overline{A_3})^{-1}} (A_1 \otimes \overline{A_4}) \right], \quad (1.11)$$

where  $P_1 = \sum_{l,m} |l \otimes l\rangle \langle m \otimes m|$ . The  $n = 4$  case is particularly relevant as it can be applied to recoverability questions in quantum information theory.<sup>16</sup> We explore the potential usefulness of the new formulation (1.11) in this context at the end of the Introduction.

- (v) Our proof of Theorem 1.1 relies on the  $n$ -matrix inequality (1.4) as input. Specifically, Theorem 1.1 is proved by showing that the right-hand side of inequality (1.4) can be rewritten as the right-hand side of (1.10), up to a permutation of  $A_k$  (Lemma 2.1).

The maximally entangled states arise via the following elementary property: For any two self-adjoint matrices  $X$  and  $Y$  on  $\mathcal{H}^{\otimes m}$ , we have

$$\mathrm{Tr}_{\mathcal{H}^{\otimes m}}[XY] = \mathrm{Tr}_{\mathcal{H}^{\otimes 2m}}[P_m(X \otimes Y^T)] = \mathrm{Tr}_{\mathcal{H}^{\otimes 2m}}[P_m(X \otimes \bar{Y})]. \quad (1.12)$$

The basic idea of the proof of Lemma 2.1 is to use this identity iteratively on the right-hand side of (1.4), splitting in half the set of matrices  $A_2, \dots, A_{n-1}$  that come with a complex power at every iteration step. The iteration stops after  $n'$  steps, when one obtains an expression that can be rewritten via (1.6). At every step, the Hilbert space is tensored with itself as in (1.12), and this is what produces the projections  $P_{2^j}$  in (1.10).

- (vi) By (1.3), we can rewrite the right-hand side of (1.10) in terms of the Fréchet derivative of a matrix logarithm. Then, (1.10) reads

$$\begin{aligned} & \mathrm{Tr}_{\mathcal{H}} \left[ \exp \left( \sum_{k=1}^n \log A_k \right) \right] \\ & \leq \frac{d}{dr} \Big|_{r=0} \mathrm{Tr}_{\mathcal{H}^{\otimes 2n'}} \left[ P_{n'-1} \exp \left( -\log \mathcal{A} + \log \left( \mathcal{A} + rA_1 \otimes \bar{A}_n \otimes \bigotimes_{j=0}^{n'-2} P_{2^j} \right) \right) \right], \end{aligned}$$

where we abbreviated  $\mathcal{A} := \bigotimes_{k=2}^{n-1} C^{\alpha_k} A_k^{-1} C^{\alpha_k} \otimes I_{\mathcal{H}}^{\otimes \rho}$ . By using homogeneity as in Lemma 5 of Ref. 12, we see that this inequality is related to the concavity of the homogeneous trace function

$$K \mapsto \mathrm{Tr}_{\mathcal{H}^{\otimes 2n'}} [P_{n'-1} \exp(-\log \mathcal{A} + \log K)]$$

near the matrix  $K = \mathcal{A}$ . In particular, a direct proof of the concavity of this trace function would yield a new, direct proof of Theorem 1.1 that does not rely on (1.4).

## B. A related inequality

Notice that the right-hand side in (1.10) is of the general form  $\mathrm{Tr}[PT_X(Y)]$  for any  $n$ , where  $P$  is a non-negative definite matrix and  $X$  and  $Y$  are appropriate positive definite matrices. Therefore, for any  $n$ , it can be estimated from below via Lieb's original triple matrix inequality (1.7), up to an approximation argument to make  $P$  strictly positive definite.

Here we discuss what one gets from this approach: An apparently weaker inequality than (1.10) which we state in Proposition 1.3 below, in the case  $n = 4$ . (The result generalizes to higher  $n$ .)

*Proposition 1.3.* Let  $A_1, A_2, A_3$ , and  $A_4$  be positive definite matrices on a  $d$ -dimensional Hilbert space  $\mathcal{H}$ . Then we have

$$d \exp \left( \frac{1}{d} \mathrm{Tr}[\log A_1 + \log A_2 + \log A_3 + \log A_4] \right) \leq \mathrm{Tr}_{\mathcal{H}^{\otimes 2}} \left[ P_1 T_{(A_2 \otimes \bar{A}_3)^{-1}} (A_1 \otimes \bar{A}_4) \right]. \quad (1.13)$$

The statement and proof of Proposition 1.3 generalize to other values of  $n > 4$ .

*Remark 1.4.* The right-hand side in (1.13) can alternatively be expressed in terms of complex matrix powers by using (2.2).

## C. Discussion

The main purpose of Theorem 1.1 is to provide an  $n$ -matrix generalization of Lieb's original triple matrix inequality (1.2).

Given that this generalization is equivalent to (and derived from) the  $n$ -matrix inequality of Sutter, Berta, and Tomamichel (1.4), one may ask if the alternative formulation (1.10) affords any advantages as compared to the original (1.4).

The clear advantage of (1.10) is the absence of complex matrix powers. We find this aesthetically pleasing since one is bounding a manifestly real-valued quantity  $\text{Tr}[\exp(\sum_k \log A_k)]$ . More importantly, the complex matrix powers are replaced by *resolvents* (e.g.,  $A_2$  appears as  $(A_2^{-1} \otimes B + \tau)^{-1}$ , where  $\tau > 0$  and  $B$  is independent of  $A_2$ ), and these have several *technical advantages* for analysis, which we review below.

A disadvantage of (1.10) is the tensor product structure that increases in complexity with  $n$ . However, for small  $n$ , the tensor product structure is still quite manageable; see, e.g., the case  $n = 4$  in (1.11).

We recall two of the technical advantages of general resolvents:

1. Resolvents are well-suited for performing *perturbation theory* since the perturbation series has an explicit and simple form at every finite order. In particular, the Fréchet derivative of a resolvent can be computed explicitly.
2. The *commutator* with a resolvent can be written in a simple form.

The particular resolvents that appear in Theorem 1.1 possess an additional advantage: Since they are of the form  $(X + \tau)^{-1}$  with  $X$  positive definite and  $\tau > 0$ , they are themselves positive definite matrices and can be estimated straightforwardly. For example, we have the operator inequalities

$$0 \leq \frac{1}{X + \tau} \leq \min\{\lambda_{\min}(X)^{-1}, \tau^{-1}\},$$

where  $\lambda_{\min}(X) > 0$  denotes the smallest eigenvalue of  $X$ .

Let us explain the potential usefulness of point 2 from above further. In short, it allows us to compare the  $n$ -matrix inequality (1.4) to variants in which the matrices  $A_k$  have been reordered, e.g., to remove the complex powers. More precisely, one can bound the difference between the right-hand side in (1.4) and related expressions without complex powers *explicitly in terms of the commutators between the different  $A_k$* . As an example for this, we take  $n = 3$  for simplicity and rewrite the difference

$$\begin{aligned} & A_1 A_2 - \int_{\mathbb{R}} A_2^{\frac{1+i\tau}{2}} A_1 A_2^{\frac{1-i\tau}{2}} \beta(\tau) d\tau \\ &= \int_0^\infty \left( A_1 \left( \frac{1}{A_2^{-1} + \tau} \right)^2 - \frac{1}{A_2^{-1} + \tau} A_1 \frac{1}{A_2^{-1} + \tau} \right) d\tau \\ &= \int_0^\infty \left[ A_1, \frac{1}{A_2^{-1} + \tau} \right] \frac{1}{A_2^{-1} + \tau} d\tau \\ &= \int_0^\infty \frac{1}{A_2^{-1} + \tau} A_2^{-1} [A_1, A_2] A_2^{-1} \left( \frac{1}{A_2^{-1} + \tau} \right)^2 d\tau. \end{aligned}$$

Note that the last expression is linear in the commutator  $[A_1, A_2]$ .

The above-mentioned technical advantages of (1.10) are only meaningful in the context of *applications* of the  $n$ -matrix inequalities (1.4) and (1.10). Here we discuss as an example for the application of the four-matrix inequality to *quantum information theory*.

Motivated by previous results in this direction,<sup>2,5,17,18,21</sup> and especially Ref. 10, the authors of Ref. 16 used their new four-matrix inequality to derive a strengthened data processing inequality, with a remainder term that involves the measured relative entropy of an appropriately “rotated” Petz recovery channel, called  $\mathcal{R}$  below.

From these papers, the rotated Petz recovery channel has emerged as a natural variant of the original Petz recovery channel.<sup>13</sup>

To see how Theorem 1.1 can be useful in this context, we simply replace the application of the four-matrix inequality (1.4) in Eq. (54) of Ref. 16 with the new formulation (1.11). [Equivalently, we can apply (2.2) to the result of Theorem 4.1 in Ref. 16.] In this way, we find the following alternative to the strengthened data processing inequality from Ref. 16. Using the notation from Ref. 16, we have

$$\begin{aligned}
& D(\rho_{AB} \parallel \sigma_{AB}) - D(\rho_A \parallel \sigma_A) \\
& \geq D_{\mathbb{M}}(\rho_{AB} \parallel \mathcal{R}_{\sigma_{AB}, \text{Tr}_B}(\rho_A)) \\
& = \sup_{\omega_{AB} > 0} \left\{ \text{Tr}_{\mathcal{H}_{AB}} [\rho_{AB} \log \omega_{AB}] + 1 - \text{Tr}_{\mathcal{H}_{AB}^{\otimes 2}} \left[ P_1 T_{\sigma_A \otimes I_B \otimes \overline{\sigma_{AB}^{-1}}} (\rho_A \otimes I_B \otimes \overline{\omega_{AB}}) \right] \right\}. \quad (1.14)
\end{aligned}$$

We mention that (1.14) can also be derived from the quantum Bayes' rule, Eq. (8) in the recent preprint,<sup>3</sup> via (1.3) and (1.12).

The equality in (1.14) provides a tool to perform analysis on the performance of the rotated Petz channel  $\mathcal{R}_{\sigma_{AB}, \text{Tr}_B}(\rho_A)$  by using (1.3) and the properties of resolvents mentioned above. For example, one can investigate

1. the change of the performance of the rotated Petz recovery channel under a perturbation of the involved states, say  $\sigma_{AB} \rightarrow \tilde{\sigma}_{AB}$ , and
2. the difference between the rotated Petz recovery channel and the original Petz recovery channel, i.e., the effect of removing the complex powers.

Note that after performing analysis with resolvents, one can return to the formulation used in Ref. 16 via (2.2) and more generally Lemma 2.1.

To summarize this example, we hope that the alternative way to write the strengthened data processing inequality (1.14) can be of value in analyzing the rotated Petz recovery channel, which has recently emerged as an important object in quantum information theory.

## II. PROOFS

### A. Proof of the main result, theorem 1.1

The proof of Theorem 1.1 is based on the following key lemma. For its statement, we recall the notation introduced before and within Theorem 1.1.

*Lemma 2.1.* Let  $n \geq 3$ . There exists a permutation  $\pi$  of  $\{2, \dots, n-1\}$  such that the following holds. Let  $X_1, \dots, X_n$  be non-negative definite matrices on a finite-dimensional Hilbert space  $\mathcal{H}$ . Then we have

$$\begin{aligned}
& \text{Tr}_{\mathcal{H}} \left[ X_n X_{n-1}^{\frac{1+i}{2}} \dots X_2^{\frac{1+i}{2}} X_1 X_2^{\frac{1-i}{2}} \dots X_{n-1}^{\frac{1-i}{2}} \right] \\
& = \text{Tr}_{\mathcal{H}^{\otimes 2n'}} \left[ P_{2^{n'-1}} \left( \left( \bigotimes_{k=2}^{n-1} C^{\alpha_k} X_{\pi(k)} C^{\alpha_k} \right)^{\frac{1+i}{2}} \otimes I_{\mathcal{H}^{\otimes \rho}} \right) \left( X_1 \otimes \overline{X_n} \otimes \bigotimes_{j=0}^{n'-2} P_{2^j} \right) \right. \\
& \quad \left. \left( \left( \bigotimes_{k=2}^{n-1} C^{\alpha_k} X_{\pi(k)} C^{\alpha_k} \right)^{\frac{1-i}{2}} \otimes I_{\mathcal{H}^{\otimes \rho}} \right) \right] \quad (2.1)
\end{aligned}$$

for every  $t \in \mathbb{R}$ .

*Remark 2.2.* The permutation  $\pi$  is explicit but of no further interest to us. Here we just mention that  $\pi = \text{id}$  when  $n = 4$ , i.e.,

$$\int_{\mathbb{R}} \text{Tr} [A_4 A_3^{\frac{1+i}{2}} A_2^{\frac{1+i}{2}} A_1 A_2^{\frac{1-i}{2}} A_3^{\frac{1-i}{2}}] \beta(t) dt = \text{Tr}_{\mathcal{H}^{\otimes 2}} \left[ P_1 T_{(A_2 \otimes \overline{A_3})^{-1}} (A_1 \otimes \overline{A_4}) \right], \quad (2.2)$$

and  $\pi \neq \text{id}$  when  $n > 4$ .

We will prove Lemma 2.1 below, by induction. Assuming that it holds, we can give the following:

*Proof of Theorem 1.1.* Fix  $n \geq 3$  and let  $\pi$  be the corresponding permutation from Lemma 2.1. The idea is to apply inequality (1.4) from Ref. 16 to a reordered collection of  $\{A_k\}$ , namely to

$$X_k := \begin{cases} A_k, & \text{if } k = 1 \text{ or } k = n, \\ A_{\pi^{-1}(k)}, & \text{if } 2 \leq k \leq n-1. \end{cases}$$

We first apply (1.4) and then Lemma 2.1 to find

$$\begin{aligned} & \operatorname{Tr} \left[ \exp \left( \sum_{k=1}^n \log A_k \right) \right] = \operatorname{Tr} \left[ \exp \left( \sum_{k=1}^n \log X_k \right) \right] \\ & \leq \int_{\mathbb{R}} \operatorname{Tr} [X_n X_{n-1}^{\frac{1+it}{2}} \cdots X_2^{\frac{1+it}{2}} X_1 X_2^{\frac{1-it}{2}} \cdots X_{n-1}^{\frac{1-it}{2}}] \beta(t) dt \\ & = \int_{\mathbb{R}} \operatorname{Tr}_{\mathcal{H}^{\otimes 2n'}} \left[ P_{2^{n'-1}} \left( \bigotimes_{k=2}^{n-1} C^{\alpha_k} A_k C^{\alpha_k} \right)^{\frac{1+it}{2}} \otimes I_{\mathcal{H}^{\otimes \rho}} \right] \left( A_1 \otimes \overline{A_n} \otimes \bigotimes_{j=0}^{n'-2} P_{2^j} \right) \\ & \quad \left( \left( \bigotimes_{k=2}^{n-1} C^{\alpha_k} A_k C^{\alpha_k} \right)^{\frac{1-it}{2}} \otimes I_{\mathcal{H}^{\otimes \rho}} \right) \beta(t) dt. \end{aligned}$$

In the last step, we used the equality  $X_{\pi(k)} = A_k$  for  $2 \leq k \leq n-1$ . Theorem 1.1 now follows from the general fact that  $Y^{\frac{1+it}{2}} \otimes I = (Y \otimes I)^{\frac{1+it}{2}}$ , where  $Y$  is any positive matrix, and the equality (1.7).  $\square$

It remains to give the following:

*Proof of Lemma 2.1.* Step 1. We first show that it suffices to prove the claim for all  $n$  of the form  $n = 2^N + 2$  for some  $N \geq 0$  (i.e., those with  $\rho = 0$ ).

Suppose that the claim holds for all  $n$  of the form  $n = 2^N + 2$ . Consider any  $n_0$  not of this form, i.e., any  $n_0$  satisfying  $2^{N_0} < n_0 - 2 < 2^{N_0+1}$  for some  $N_0 \geq 0$ . Since the claim holds for  $n = 2^{N_0+1} + 2$ , we get a corresponding permutation  $\pi$  of  $\{2, \dots, 2^{N_0+1} + 1\}$ . We choose matrices  $\{Y_k\}_{k=1}^{2^{N_0+1}+2}$  as follows:

$$Y_k := \begin{cases} X_1, & \text{if } k = 1, \\ X_{n_0}, & \text{if } k = 2^{N_0+1} + 2, \\ X_k, & \text{if } 2 \leq \pi(k) \leq n_0 - 1, \\ I_{\mathcal{H}}, & \text{if } n_0 + 1 \leq \pi(k) \leq 2^{N_0+1} + 1. \end{cases}$$

By assumption, we have the analogue of (2.1) for  $\{Y_k\}$ . Using  $n' = N_0 + 1$ , this means

$$\begin{aligned} & \operatorname{Tr} [Y_{2^{N_0+1}+2} Y_{2^{N_0+1}+1}^{\frac{1+it}{2}} \cdots Y_2^{\frac{1+it}{2}} Y_1 Y_2^{\frac{1-it}{2}} \cdots Y_{2^{N_0+1}+1}^{\frac{1-it}{2}}] \\ & = \operatorname{Tr}_{\mathcal{H}^{\otimes 2^{N_0+1}}} \left[ P_{2^{N_0}} \left( \bigotimes_{k=2}^{2^{N_0+1}+1} C^{\alpha_k} Y_{\pi(k)} C^{\alpha_k} \right)^{\frac{1+it}{2}} \left( Y_1 \otimes \overline{Y_{2^{N_0+2}}} \otimes \bigotimes_{j=0}^{N_0-1} P_{2^j} \right) \right. \\ & \quad \left. \left( \bigotimes_{k=2}^{2^{N_0+1}+1} C^{\alpha_k} Y_{\pi(k)} C^{\alpha_k} \right)^{\frac{1-it}{2}} \right]. \end{aligned}$$

Thanks to the definitions of  $\{Y_k\}$  and  $\rho = 2^{N_0+1} - n_0 + 2$ , this is simply the claimed Eq. (2.1) for  $\{X_k\}_{k=1}^{n_0}$  and an appropriate permutation. This concludes step 1.

Step 2. It remains to prove Lemma 2.1 for all  $n$  of the form  $n = 2^N + 2$ . To this end, we perform an induction in  $N$ .

The *induction base* corresponds to  $N = 0$  or  $n = 3$ . It follows by combining the identity (1.6), which was already proved in Ref. 16 with the equality in Remark 1.2 (iii).

We come to the *induction step*. Fix  $N \geq 0$ . We assume that Lemma 2.1 holds for  $n = 2^N + 2$ , on all finite-dimensional Hilbert spaces and for all choices of orthonormal bases defining a maximally entangled state. We fix an arbitrary finite-dimensional Hilbert space  $\mathcal{H}$  and an orthonormal basis  $\{|l\rangle\}$ . We want to show that Lemma 2.1 holds for  $n_0 = 2^{N+1} + 2$  on this Hilbert space.



By cyclicity of the trace and the key property (1.12) of maximally entangled states, we have for every  $t \in \mathbb{R}$ ,

$$\begin{aligned} & \text{Tr}_{\mathcal{H}}[X_{n_0} X_{n_0-1}^{\frac{1+i}{2}} \dots X_2^{\frac{1+i}{2}} X_1 X_2^{\frac{1-i}{2}} \dots X_{n_0-1}^{\frac{1-i}{2}}] \\ &= \text{Tr}_{\mathcal{H}^{\otimes 2}} \left[ P_1 \left( X_{\frac{n_0}{2}}^{\frac{1+i}{2}} \dots X_2^{\frac{1+i}{2}} X_1 X_2^{\frac{1-i}{2}} \dots X_{\frac{n_0}{2}}^{\frac{1-i}{2}} \otimes \overline{X_{\frac{n_0}{2}+1}^{\frac{1-i}{2}} \dots X_{n_0-1}^{\frac{1-i}{2}} X_{n_0} X_{n_0-1}^{\frac{1+i}{2}} \dots X_{\frac{n_0}{2}+1}^{\frac{1+i}{2}}} \right) \right]. \end{aligned}$$

Using the fact that  $\overline{X_k^{\frac{1+i}{2}}} = \overline{X_k}^{\frac{1-i}{2}}$ , as well as some general rules for tensor products of matrices, we can rewrite this expression as follows:

$$\begin{aligned} & \text{Tr}_{\mathcal{H}^{\otimes 2}} \left[ P_1 \left( X_{\frac{n_0}{2}}^{\frac{1+i}{2}} \dots X_2^{\frac{1+i}{2}} X_1 X_2^{\frac{1-i}{2}} \dots X_{\frac{n_0}{2}}^{\frac{1-i}{2}} \otimes \overline{X_{\frac{n_0}{2}+1}^{\frac{1-i}{2}} \dots X_{n_0-1}^{\frac{1-i}{2}} X_{n_0} X_{n_0-1}^{\frac{1+i}{2}} \dots X_{\frac{n_0}{2}+1}^{\frac{1+i}{2}}} \right) \right] \\ &= \text{Tr}_{\mathcal{H}^{\otimes 2}} \left[ P_1 \left( X_{\frac{n_0}{2}}^{\frac{1+i}{2}} \otimes \overline{X_{\frac{n_0}{2}+1}^{\frac{1+i}{2}}} \right) \dots \left( X_2^{\frac{1+i}{2}} \otimes \overline{X_{n_0-1}^{\frac{1+i}{2}}} \right) \left( X_1 \otimes \overline{X_{n_0}} \right) \right. \\ & \quad \left. \left( X_2^{\frac{1-i}{2}} \otimes \overline{X_{n_0-1}^{\frac{1-i}{2}}} \right) \dots \left( X_{\frac{n_0}{2}}^{\frac{1-i}{2}} \otimes \overline{X_{\frac{n_0}{2}+1}^{\frac{1-i}{2}}} \right) \right] \\ &= \text{Tr}_{\mathcal{H}^{\otimes 2}} \left[ P_1 \left( X_{\frac{n_0}{2}} \otimes \overline{X_{\frac{n_0}{2}+1}} \right)^{\frac{1+i}{2}} \dots \left( X_2 \otimes \overline{X_{n_0-1}} \right)^{\frac{1+i}{2}} \left( X_1 \otimes \overline{X_{n_0}} \right) \right. \\ & \quad \left. \left( X_2 \otimes \overline{X_{n_0-1}} \right)^{\frac{1-i}{2}} \dots \left( X_{\frac{n_0}{2}} \otimes \overline{X_{\frac{n_0}{2}+1}} \right)^{\frac{1-i}{2}} \right]. \end{aligned}$$

Now we are in a position to apply the *induction hypothesis* for  $n = 2^N + 2 = n_0/2 + 1$  on the finite-dimensional Hilbert space  $\mathcal{H}^{\otimes 2}$  equipped with the orthonormal basis  $\{|l \otimes l'\rangle\}$ . We define

$$Z_k := X_k \otimes \overline{X_{n_0+1-k}}, \quad 1 \leq k \leq n_0/2. \quad (2.3)$$

We set  $n'_0 := \lceil \log_2(n_0 - 2) \rceil = N + 1$ . The induction hypothesis yields a permutation, call it  $\sigma$ , of  $\{2, \dots, n_0/2\}$  such that

$$\begin{aligned} & \text{Tr}_{\mathcal{H}}[X_{n_0} X_{n_0-1}^{\frac{1+i}{2}} \dots X_2^{\frac{1+i}{2}} X_1 X_2^{\frac{1-i}{2}} \dots X_{n_0-1}^{\frac{1-i}{2}}] \\ &= \text{Tr}_{\mathcal{H}^{\otimes 2}} \left[ P_1 Z_{\frac{n_0}{2}}^{\frac{1+i}{2}} \dots Z_2^{\frac{1+i}{2}} Z_1 Z_2^{\frac{1-i}{2}} \dots Z_{\frac{n_0}{2}}^{\frac{1-i}{2}} \right] \\ &= \text{Tr}_{\mathcal{H}^{\otimes 2^{n'_0}}} \left[ P_{2^{n'_0-1}} \left( \bigotimes_{l=2}^{n_0/2} C^{\alpha_l} Z_{\sigma(l)} C^{\alpha_l} \right)^{\frac{1+i}{2}} \left( Z_1 \otimes P_1 \otimes \bigotimes_{j=1}^{n'_0-2} P_{2^j} \right) \right. \\ & \quad \left. \left( \bigotimes_{l=2}^{n_0/2} C^{\alpha_l} Z_{\sigma(l)} C^{\alpha_l} \right)^{\frac{1-i}{2}} \right]. \end{aligned} \quad (2.4)$$

To arrive at the last line, we used that we can identify (non-normalized) projectors  $P_{2^j}(\{|l \otimes l'\rangle\}) = P_{2^j+1}(\{|l\rangle\})$  (where we added the dependence of the projector on the underlying basis in parentheses for emphasis). This observation allows us to replace all the projectors arising via the induction hypothesis, which are associated with the basis  $\{|l \otimes l'\rangle\}$  so that (2.4) only features projectors associated with the basis  $\{|l\rangle\}$ . We also used  $n' = N = n'_0 - 1$  and the fact that  $\overline{P_1} = P_1$ , which holds because complex conjugation is defined in the eigenbasis of  $P_1$ .

Given (2.4), the claim (2.1) will follow once we prove the following:

*Lemma 2.3. There exists a permutation  $\pi$  of  $\{2, \dots, n_0 - 1\}$  such that*

$$\bigotimes_{l=2}^{n_0/2} C^{\alpha_l} Z_{\sigma(l)} C^{\alpha_l} = \bigotimes_{k=2}^{n_0-1} C^{\alpha_k} X_{\pi(k)} C^{\alpha_k}. \quad (2.5)$$

We now prove this lemma. For convenience, we rewrite the claim (2.5) in terms of the shifted Thue-Morse sequence  $\beta_j := \alpha_{j+2}$  ( $j \geq 0$ ) and the shifted permutations  $\tilde{\sigma}$  and  $\tilde{\pi}$  defined by

$$\tilde{\sigma}(j) := \sigma(j+2), \quad \tilde{\pi}(j) := \pi(j+2).$$

Using these definitions and recalling (2.3), we can write the claim (2.5) as

$$\bigotimes_{j=0}^{n_0/2-2} C^{\beta_j} (X_{\tilde{\sigma}(j)} \otimes \overline{X_{n_0+1-\tilde{\sigma}(j)}}) C^{\beta_j} = \bigotimes_{m=0}^{n_0-3} C^{\beta_m} X_{\tilde{\pi}(m)} C^{\beta_m}, \quad (2.6)$$

for an appropriately chosen bijection  $\tilde{\pi} : \{0, \dots, n_0 - 3\} \rightarrow \{2, \dots, n_0 - 1\}$ .

We recall that  $\tilde{\sigma}$  is a bijection onto  $\{2, \dots, n_0/2\}$ . We see that each  $X_k$  from the set  $\{X_k\}_{k=2}^{n_0-1}$  appears precisely once on both sides of (2.6). Hence, we can simply read off the appropriate definition of  $\tilde{\pi}$  (and therefore of  $\pi$ ) from (2.6). That is, we take

$$\tilde{\pi}(m) := \begin{cases} \tilde{\sigma}\left(\frac{m}{2}\right), & \text{if } m \text{ is even,} \\ n_0 + 1 - \tilde{\sigma}\left(\frac{m-1}{2}\right), & \text{if } m \text{ is odd.} \end{cases}$$

With this definition, the claim (2.6) is equivalent to

$$\bigotimes_{j=0}^{n_0/2-2} C^{\beta_j} (X_{\tilde{\sigma}(j)} \otimes \overline{X_{n_0+1-\tilde{\sigma}(j)}}) C^{\beta_j} = \bigotimes_{j=0}^{n_0/2-2} \left( C^{\beta_{2j}} X_{\tilde{\sigma}(j)} C^{\beta_{2j}} \otimes C^{\beta_{2j+1}} X_{n_0+1-\tilde{\sigma}(j)} C^{\beta_{2j+1}} \right).$$

We can now conclude this equality from the relation (1.8), observing in particular that it yields

$$C^{\beta_{2j+1}} X_{n_0+1-\tilde{\sigma}(j)} C^{\beta_{2j+1}} = C^{1-\beta_j} X_{n_0+1-\tilde{\sigma}(j)} C^{1-\beta_j} = C^{\beta_j} \overline{X_{n_0+1-\tilde{\sigma}(j)}} C^{\beta_j}.$$

This proves Lemma 2.3 and thus finishes step 2 of the proof of Lemma 2.1.  $\square$

## B. Proof of proposition 1.3

We begin from the right-hand side of (1.13). We introduce the normalized projectors

$$\tilde{P}_1 := d^{-1} P_1, \quad \tilde{P}_1^\perp := I_{\mathcal{H}^{\otimes 2}} - \tilde{P}_1,$$

which are true projection operators in the sense that their eigenvalues are zero and one. We approximate  $\tilde{P}_1$  by the strictly positive definite matrix

$$\tilde{P}_\delta := \tilde{P}_1 + \delta I_{\mathcal{H}^{\otimes 2}}, \quad 0 < \delta < 1.$$

After implementing the approximation, we can apply Lieb's triple matrix inequality (1.2) to get

$$\begin{aligned} & \text{Tr}_{\mathcal{H}^{\otimes 2}} \left[ P_1 T_{(A_2 \otimes \overline{A_3})^{-1}} (A_1 \otimes \overline{A_4}) \right] \\ &= d \text{Tr}_{\mathcal{H}^{\otimes 2}} \left[ \tilde{P}_1 T_{(A_2 \otimes \overline{A_3})^{-1}} (A_1 \otimes \overline{A_4}) \right] = d \lim_{\delta \rightarrow 0} \text{Tr}_{\mathcal{H}^{\otimes 2}} \left[ \tilde{P}_\delta T_{(A_2 \otimes \overline{A_3})^{-1}} (A_1 \otimes \overline{A_4}) \right] \\ &\geq d \liminf_{\delta \rightarrow 0} \text{Tr}_{\mathcal{H}^{\otimes 2}} \left[ \exp \left( \log \tilde{P}_\delta + \log(A_2 \otimes \overline{A_3}) + \log(A_1 \otimes \overline{A_4}) \right) \right]. \end{aligned}$$

By the spectral theorem,

$$\log \tilde{P}_\delta = \log(1 + \delta) \tilde{P}_1 + \log(\delta) \tilde{P}_1^\perp \geq \log(\delta) \tilde{P}_1^\perp.$$

Thus, after changing variables to  $t := -\log \delta$ , we get

$$\begin{aligned} & d \liminf_{\delta \rightarrow 0} \text{Tr} \left[ \exp \left( \log \tilde{P}_\delta + \log(A_2 \otimes \overline{A_3}) + \log(A_1 \otimes \overline{A_4}) \right) \right] \\ &\geq d \liminf_{t \rightarrow \infty} \text{Tr} \left[ \exp \left( -t \tilde{P}_1^\perp + \log(A_2 \otimes \overline{A_3}) + \log(A_1 \otimes \overline{A_4}) \right) \right]. \end{aligned} \quad (2.7)$$

The right-hand side can be computed explicitly via the following lemma. The lemma can be seen as an asymptotic (and therefore much simpler) version of Stahl's formula, derived in his proof of the BMV (Bessis, Moussa, and Villani) conjecture.<sup>14</sup>

*Lemma 2.4.* Let  $A$  be a self-adjoint matrix and let  $P$  be a projection so that  $\ker P = \text{span}\{v\}$  for a normalized vector  $v$ . Then

$$\lim_{t \rightarrow \infty} \text{Tr}[\exp(A - tP)] = \exp \langle v, Av \rangle.$$

Before we prove the lemma, we observe that it yields Proposition 1.3. Indeed, we apply the lemma to the right-hand side in (2.7) and note that  $\ker \tilde{P}_1^\perp = \text{span}\{|\Omega_1\rangle\}$ . Since  $\|\Omega_1\| = \sqrt{d}$ , we obtain

$$\begin{aligned} & \text{Tr}_{\mathcal{H}^{\otimes 2}} \left[ P_1 T_{(A_2 \otimes \overline{A_3})^{-1}} (A_1 \otimes \overline{A_4}) \right] \\ & \geq d \liminf_{t \rightarrow \infty} \text{Tr}[\exp(-t\tilde{P}_1^\perp + \log(A_2 \otimes \overline{A_3}) + \log(A_1 \otimes \overline{A_4}))] \\ & = d \exp \left( \frac{1}{d} \langle \Omega_1, (\log(A_2 \otimes \overline{A_3}) + \log(A_1 \otimes \overline{A_4})) \Omega_1 \rangle \right) \\ & = d \exp \left( \frac{1}{d} \text{Tr}[\log A_1 + \log A_2 + \log A_3 + \log A_4] \right). \end{aligned}$$

In the last step, we used the elementary property (1.12) of the maximally entangled state  $\Omega_1$ . This proves Proposition 1.3, assuming the validity of Lemma 2.4.

It remains to give the following:

*Proof of Lemma 2.4.* Fix  $m \geq 1$ , a self-adjoint matrix  $A$  and a projector  $P$  acting on  $\mathbb{C}^m$  so that  $\ker P = \text{span}\{v\}$  for a normalized vector  $v$ . We are interested in the eigenvalues of  $A - tP$  for large  $t > 0$ , call them  $\lambda_1(t), \dots, \lambda_m(t)$ . We rewrite the matrix as

$$A - tP = t(-P + t^{-1}A),$$

and we let  $\mu_1(t), \dots, \mu_m(t)$  denote the eigenvalues of  $(-P + t^{-1}A)$  so that  $\lambda_j(t) = t\mu_j(t)$ . This places the problem in the setting of standard analytic perturbation theory, with  $t^{-1}$  being the small parameter. Using the results from chapter 7 (§ 2), of Kato's book,<sup>11</sup> it is straightforward to compute the eigenvalues  $\mu_j(t)$  to lowest order in  $t^{-1}$ . (Here we use our assumption that  $\ker P$  is one-dimensional, as it essentially reduces the problem to non-degenerate perturbation theory.)

Using  $\lambda_j(t) = t\mu_j(t)$ , we obtain, up to relabeling of the eigenvalues,

$$\lambda_1(t) = \langle v, Av \rangle + o(1), \quad \lambda_j(t) = -t + o(t), \quad (j = 2, \dots, m).$$

Therefore, as  $t \rightarrow \infty$ ,

$$\text{Tr}[\exp(A - tP)] = \sum_{j=1}^m e^{\lambda_j(t)} = e^{\lambda_1(t)} + o(1) = e^{\langle v, Av \rangle} + o(1).$$

This proves Lemma 2.4 and thus also Proposition 1.3.  $\square$

## ACKNOWLEDGMENTS

The author is grateful to Elliott H. Lieb for raising the question addressed in this paper. It is a pleasure to thank him as well as Mario Berta, Eric Carlen, Rupert L. Frank, and David Sutter for helpful remarks.

<sup>1</sup> Bergh, J. and Löfström, J., *Interpolation Spaces* (Springer, 1976).

<sup>2</sup> Berta, M., Lemm, M., and Wilde, M. M., "Monotonicity of quantum relative entropy and recoverability," *Quantum Inf. Comput.* **15**, 1333 (2015).

<sup>3</sup> Cotler, J., Hayden, P., Salton, G., Swingle, B., Walter, M., "Entanglement wedge reconstruction via universal recovery channels," preprint [arXiv:1704.05839](https://arxiv.org/abs/1704.05839).

<sup>4</sup> Dupuis, F. and Wilde, M. M., "Swiveled Rényi entropies," *Quantum Inf. Process.* **15**(3), 1309–1345 (2016).

<sup>5</sup> Fawzi, O. and Renner, R., "Quantum conditional mutual information and approximate Markov chains," *Commun. Math. Phys.* **340**, 575 (2015).

<sup>6</sup> Golden, S., "Lower bounds for the Helmholtz function," *Phys. Rev.* **137**, B1127–B1128 (1965).

<sup>7</sup> Hansen, F., "Multivariate extensions of the Golden-Thompson inequality," *Ann. Funct. Anal.* **6**(4), 301–310 (2015).

<sup>8</sup> Hiai, F., Köenig, R., and Tomamichel, M., "Generalized log-majorization and multivariate trace inequalities," *Ann. Henri Poincaré* **18**(7), 2499–2521 (2017).

<sup>9</sup> Hirschman, I. I., "A convexity theorem for certain groups of transformations," *J. Anal. Math.* **2**(2), 209–218 (1952).

- <sup>10</sup> Junge, M., Renner, R., Sutter, D., Wilde, M. M., and Winter, A., “Universal recovery from a decrease of quantum relative entropy,” e-print [arXiv:1509.07127](https://arxiv.org/abs/1509.07127).
- <sup>11</sup> Kato, T., *Perturbation Theory for Linear Operators* (Springer, 1980).
- <sup>12</sup> Lieb, E. H., “Convex trace functions and the Wigner-Yanase-Dyson conjecture,” *Adv. Math.* **11**, 267–288 (1973).
- <sup>13</sup> Petz, D., “Sufficiency of channels over von Neumann algebras,” *Q. J. Math.* **39**(1), 97–108 (1988).
- <sup>14</sup> Stahl, H. R., “Proof of the BMV conjecture,” *Acta Math.* **211**(2), 255–290 (2013).
- <sup>15</sup> Stein, E. M., “Interpolation of linear operators,” *Trans. Am. Math. Soc.* **83**(2), 482–492 (1956).
- <sup>16</sup> Sutter, D., Berta, M., and Tomamichel, M., “Multivariate trace inequalities,” *Commun. Math. Phys.* **352**(1), 37–58 (2017).
- <sup>17</sup> Sutter, D., Fawzi, O., and Renner, R., “Universal recovery channel for approximate Markov chains,” *Proc. R. Soc. A* **472**, 20150623 (2016).
- <sup>18</sup> Sutter, D., Tomamichel, M., and Harrow, A. W., “Strengthened monotonicity of relative entropy via pinched Petz recovery map,” *IEEE Trans. Inf. Theory* **62**(5), 2907–2913 (2016).
- <sup>19</sup> Thompson, C. J., “Inequality with applications in statistical mechanics,” *J. Math. Phys.* **6**(11), 1812–1813 (1965).
- <sup>20</sup> Wilde, M. M., *Quantum Information Theory*, 2nd ed. (Cambridge University Press, 2017).
- <sup>21</sup> Wilde, M. M., “Recoverability in quantum information theory,” *Proc. R. Soc. A* **471**, 20150338 (2015).
- <sup>22</sup> Wilde, M. M., “Monotonicity of p-norms of multiple operators via unitary swivels,” e-print [arXiv:1610.01262](https://arxiv.org/abs/1610.01262).
- <sup>23</sup> Wolff, M., Quantum Channels and Operations—Guided Tour, URL: <https://www-m5.ma.tum.de/foswiki/pub/M5/.../MichaelWolf/QChannelLecture.pdf>.